Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2016-285-RC2, 2016

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

Interactive comment on "Seasonal streamflow forecasting by conditioning climatology with precipitation indices" *by* L. Crochemore et al.

Anonymous Referee #2

Received and published: 26 July 2016

This study proposes an approach to improve short- and long-range (10-90 days) streamflow forecasts by conditioning resampled historical observations based on ECMWF System 4 forecasts. The conditioning is applied on both precipitation and streamflow records. Results are compared with historical resampled streamflow and ensemble streamflow prediction (ESP) as reference forecasts. Overall, the paper is well written and provides good assessments of different model performances. Nevertheless, I am concerned with the proposed method to improve streamflow forecasts (selection of resampled data based on GCM forecasts) as well as the results (week performance of the proposed method). Therefore, I think the paper is not ready for publication and requires major revision.

Major comments:

Printer-friendly version

Discussion paper

1) The manuscript states that (P4, L9) the aim of this study is “to generate forecasts that benefit from the reliability of climatology-based ensembles and the sharpness of System 4 precipitation forecasts.” First the proposed method does not seem to benefit from the sharpness of System 4, rather the reason for increased precision (sharpness) in the conditioned forecasts is due to the reduced ensemble size which is independent of the System 4’s degree of uncertainty. Second, the results (e.g. Figures 4-5) show that except for some marginal improvements in forecasts for short lead times (Figure 4 upper row), the proposed method degrade the performance of the reference methods (CRPSS and PITSS are negative). In several instances in the manuscript (such as P9, L17) the authors discuss the improvements to the sharpness of the forecasts using their conditioning approach while reliability and performance have declined compared to the reference methods which undermines the sharpness improvements. The authors state that “. . .the PIT diagrams at 45 days show that this decrease does not affect the overall reliability of the conditioned ensembles” This again shows that the proposed method has not been able to improve upon the conventional approaches.

2) The proposed method selects forecast ensemble members based on their closeness to some statistics (P8, L17). The procedure to choose the number of ensemble members to keep, however, is not explained. Is the number of selected runs subjectively chosen? If so a sensitivity analysis needs to be conducted.

3) The method conditions the resampled precipitation and streamflow data to GCM forecasts. However, GCM forecasts are uncertain particularly at seasonal scales. That might explain why the overall results do not show improvements compared with conventional ESP. In particular, authors need to discuss how the method will perform in regions with high topographical variations (considering that the low-resolution GCMs cannot capture the regional hydroclimatic variations). Related to this please discuss why you compare the proposed conditioning approach (based on SYS4) with results of SYS4?

4) Please clarify are the statistics (section 2.4.2) calculated for each ECMWF ensemble

member separately or for the average of the 51 ensemble runs?

5) P8, L25: “when directly selecting scenarios from past streamflow observations, the last observed streamflow is added as a conditioning criterion in the computation of the Euclidian distance.” This is problematic as the last observed (previous year’s(?)) streamflow is not a good indicator of the next year’s streamflow in particular with regard to high and low flows which are driven by several hydroclimatic factors that do not necessarily repeat at consecutive years.

6) Resampled precipitation is considered to drive the hydrologic model, however, the mean interannual potential evapotranspiration is used instead of the resampled one. Considering that PET might have a substantial role in low flow forecasts, I recommend using the resampled PET as well.

7) P12, L12: “The rankings are based on the visual evaluation of Figure 5.” Visual evaluation is not an appropriate ranking approach.

8) Results of section 3.4 are based on only one drought event for one catchment and cannot provide sufficient evidence for the overall performance of the methods.

9) P6, section 2.3.1 Please elaborate further on the differences between CRPS and PIT and how they should be interpreted when they show inconsistent results (e.g. Fig 4).

10) Multi-model averaging methods (such as simple mean, Bayesian Model Averaging (BMA) etc.) (Duan et al. 2005, Najafi et al. 2015, Raftery et al. 2005) have shown to improve short and long term hydrologic forecasts. I would suggest discussing the application of these approaches to merge the ensemble of forecasts obtained from different methods in this study.

Specific comments:

- Abstract “. . .forecasts based on GCM outputs can offer sharper ensembles. . .”: does “sharper” refer to more precise? Related to this please define “sharpness” and “relia-

bility" before using these terms, in the Introduction. - L15: ECMWF System 4: Please expand the full name. - Abstract: "The four conditioned precipitation scenarios were used as input to the GR6J hydrological model to obtain eight conditioned streamflow forecast scenarios": The statement is vague as to how four precipitation scenarios result in eight streamflow scenarios? - P2, L19: ESP is one of the streamflow forecast methods which need to be discussed here. Also please note that in ESP all historical meteorological forcings can be resampled to run the hydrological model (not just precipitation as stated in LP2, L27) - P4, L3 Statement is not clear "although the ensemble conditioned from historical streamflows, which was the. . ." - P4, L12-15: Please move to the results section. - P4, L17: Please define "discrimination" - P5, L3: Please explain how many grid cells lie within each catchment in average. How was the aggregation performed? Please also indicate the forecast starting date. - P5, L23: What do you mean by "systematically"? - P5, L31-33: What is the range of KGE values? Please show the equations for KGE and 1-bias and include their ranges. - P6, L9: Please change "The CRPS averages over the evaluation period the area between the cumulative forecast distribution. . ." to "The CRPS averages the area between the cumulative forecast distribution. . . over the evaluation period." Similarly, for L12. - P7, L3: What is the "reference"? Is it HisQ? Please define. - I suggest bringing section 2.4 before section 2.3. - Figure 2: What is the difference between SUM1-3 and SUM3 - P9, L1 "The reference forecast used to compute the skill scores is historical precipitations (i.e. climatology)": Do you mean hydrologic model simulation driven by historical precipitation? - P9, L3 "SPI forecasts issued from System 4 are reliable overall and in standard precipitation conditions" please provide a reference

Raftery, Adrian E., et al. "Using Bayesian model averaging to calibrate forecast ensembles." Monthly Weather Review 133.5 (2005): 1155-1174. Najafi, M. and Moradkhani, H. (2015). "Ensemble Combination of Seasonal Streamflow Forecasts." J. Hydrol. Eng., 10.1061/(ASCE)HE.1943-5584.0001250, 04015043. Duan, Qingyun, et al. "Multi-model ensemble hydrologic prediction using Bayesian model averaging." Advances in Water Resources 30.5 (2007): 1371-1386.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-285, 2016.

Printer-friendly version

Discussion paper